3.3 LAGS System Using Data/Instruction Grain Power Control

Makoto Ikeda[1], Taku Sogabe[1], Ken Ishii[1], Masayuki Mizuno[2], Toru Nakura[2], Koichi Nose[2], Kunihiro Asada[1]

[1]University of Tokyo, Tokyo, Japan
[2]NEC, Sagamihara, Japan

A locally asynchronous, globally synchronous (LAGS) system is developed for the optimization of power supply voltage and speed performance subject to performance requirements. The efficacy of the concept is evaluated in a 90nm CMOS CPU implementation. The LAGS system detects the cycle-by-cycle execution time, based on a completion-detection architecture using DCVSL, and controls V_{dd} and hence speed, taking into account data and instruction dependent execution time distributions. This asynchronous operation guarantees reliable operation against dynamic PVT variations, switching operating speed dynamically according to V_{dd} and temperature variations. The dynamic voltage scaling architecture controls V_{dd} according to the number of executed instructions per unit time, for V_{dd} minimization, while sustaining the required performance.

Conventional synchronous systems suffer from PVT variations, which will threaten reliable operation in the near future given current technology trends. Dynamic voltage frequency scaling (DVFS) [1] systems have been proposed for power and performance optimization, which will also be severely affected by PVT variations when operating at low V_{dd} resulting in limitations on V_{dd} lowering.

Asynchronous systems, in ideal cases, achieve average operation time and eliminate unnecessary signal transitions for power reduction [2]. A LAGS system is developed based on a completion-detection type asynchronous system for global synchronization. Figure 3.3.1 shows the proposed dynamic voltage control scheme based on a LAGS system. The conventional DVFS system controls V_{dd} and clock frequency periodically (on the order of ms), based on a frequency-voltage (F-V) table or replica-delay to achieve the required operation speed. The F-V table and the replica-delay are designed to satisfy the maximum delay time (critical path delay) with a certain amount of margin, for reliable operation, against static and dynamic PVT variations. The scheme developed here compares target performance with the number of operations executed in a unit period counting completion pulses, for V_{dd} control. Speed-performance is guaranteed by V_{dd} control based on the average at-speed of each operation. This scheme is free from margins because the at-speed already includes the PVT variations, which results in lower V_{dd} operation compared with conventional DVFS schemes. Asynchronous systems however, suffer from global synchronization issues. DVFS systems face the same issue and usually employ FIFOs for global synchronization. The depth of the FIFOs are designed carefully to prevent FIFO-related CPU stalls such as FIFO full and FIFO empty. The same technique is employed here for global synchronization of the asynchronous system.

Figure 3.3.2 shows a comparison of cycle times between replica-delay synchronization [2] and completion-signal based synchronization techniques [3,4]. The former has an advantage of smaller area overhead, however, it cannot make use of data-dependent path-delay distributions, which results in operations with maximum delay with a certain amount of timing margin. It is also important to note that the design of the replica-delay lines are critical both for the performance of the chip and the reliable operation against the PVT variations. However, it is challenging to design replica-delay for operations at low V_{dd} in systems like replica-delay based DVFS systems. The latter technique, on the other hand, generates completion signals from the dual-rail signals, with overhead for both area and delay of the completion generation tree, and achieves operations taking into account data dependent path delay, as well as instruction dependent path delay, without any timing margins.

Figure 3.3.3 shows measurement results of instruction, data and V_{dd} dependent operation execution time distribution using a CPU based on a completion-detection type asynchronous architecture fabricated in 90nm CMOS. This figure implies a speed-performance advantage of the completion-detection type system over the replica-delay based system, due to the data dependency of execution time, as well as timing margin. The completion-detection type achieves 25% performance improvement over the replica-delay type in the case of 0.75V operation, where the average delay for ADD operations is 41.8ns and the worst-case delay is 47.0ns, which results in a cycle time of 56.4ns with 20% timing margin, as shown in Fig. 3.3.3. This results in further V_{dd} reduction with the same operating speed.

The LAGS system is implemented based on a completion-detection type asynchronous CPU using DCVSL [5]. Figure 3.3.4 shows a micrograph of the LAGS CPU designed and fabricated in 90nm CMOS. The non-pipelined CPU is based on a Z80 instruction-set compatible architecture with single clock-per-instruction operation. The gate count and area of the CPU is 6.7kgates and $450\times450\mu m^2$, respectively, along with on-chip 1kword data memory and 1kword instruction memory. The LAGS CPU is also designed with an on-chip DC-DC converter that occupies $250\times60\mu m^2$, based on an on-chip VCO and pulse width modulator for high-speed voltage control, as shown in Fig. 3.3.5. The rest of the chip has been filled with an on-chip de-coupling capacitance of 3nF. A simple V_{dd} control flow for the fabricated CPU is shown in Fig. 3.3.5, which is realized in an off-chip FPGA.

Figure 3.3.6 shows measured performance of the on-chip DC-DC converter and CPU. The DC-DC converter achieves V_{dd} control over a range of 300mV (0.68V to 0.98V) with 5b resolution for an external 1.2V power supply. It also achieves a transient time of 50ns and 85ns, respectively, for a 300mV rise and 300mV fall. CPU current consumption, I_{dcpu} varies from 8.3mA to 21.0mA when the CPU power supply, V_{ddcpu} varies from 0.81V to 1.02V, while the DC-DC converter input current I_{ddcdc} varies from 9.9mA to 22.5mA. A 1.2V power supply for the DC-DC converter, results in a 57% to 79% power efficiency, respectively, for V_{ddcpu} of 0.81 to 1.02V. The average CPU operating speed varies from 22.5ns to 11.31ns.

Figure 3.3.7 shows measured V_{dd} transients to illustrate target speed tracking performance. Measurement results show a speed tracking accuracy of 5% for 25MHz to 70MHz target operation speed. It is also shown that the average current consumption is reduced by 30% for operation at a target speed of 48MHz compared with a fixed 1V operation.

The proposed LAGS system achieves the required performance as defined by the target speed, even with static and the dynamic PVT variations. It is also noted that DCVSL has the potential to detect logic errors, which induce logic '1' and '1' on both dual-rail signals, that can be utilized for enhancing reliable operation, in addition to noise reduction achieved by V_{dd} lowering.

References:

[1] S. Akui, et al., "Dynamic Voltage and Frequency Management for a Low-Power Embedded Microprocessor," *ISSCC Dig. Tech. Papers*, pp. 64-66, Feb., 2004.
[2] Arjan Bink, "ARM996HS: The First Licensable, Clockless 32-Bit Processor Core," Session 2-1, Spring Processor Forum, May 2006.
[3] K. R. Cho, K. Okura and K. Asada, "Experimental Design of a 32-bit Fully Asynchronous Microprocessor (FAM)," *IEICE Trans. on Electronics*, vol. E77-C, no. 4, pp. 615-623, Apr. 1994.
[4] T. Nanya, et al., "TITAC-2: A 32-bit Scalable-Delay-Insensitive Microprocessor," *HOT Chips IX, Stanford*, pp.19-32, Aug., 1997.
[5] L. G. Heller, et. al., "Cascode Voltage Switch Logic: A Differential CMOS Logic Family," *ISSCC Dig. Tech. Papers*, pp. 16-17, Feb. 1984.

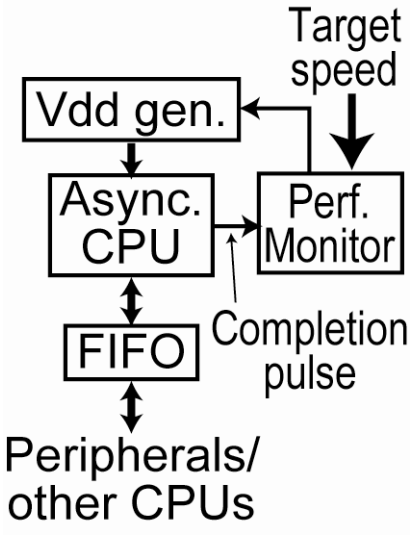

Figure 3.3.1: The LAGS system block diagram.

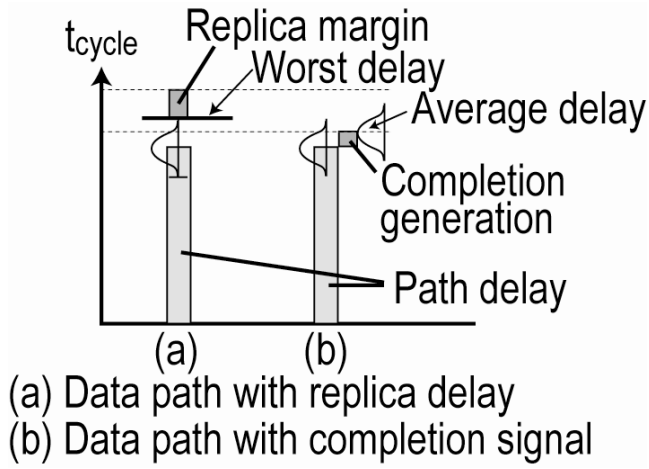

Figure 3.3.2: Comparison of cycle times.

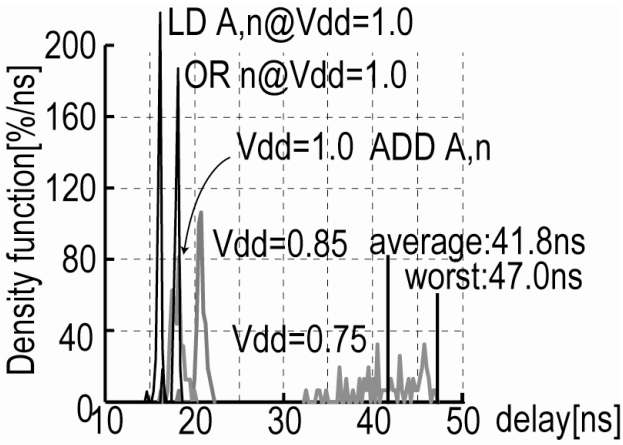

Figure 3.3.3: Measured results of instruction, data, and Vdd dependent operation execution time distribution.

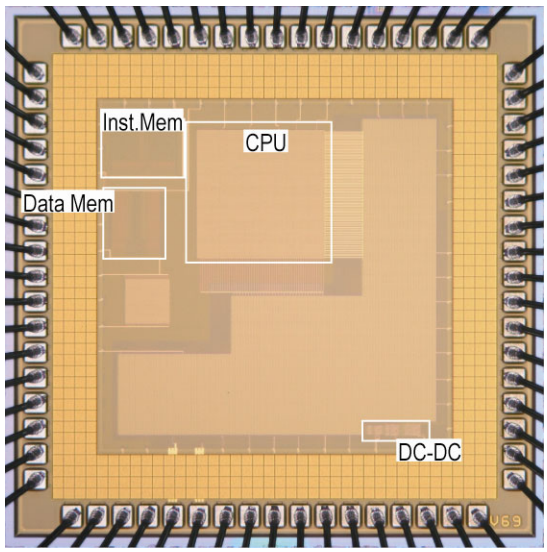

Figure 3.3.4: Chip micrograph.

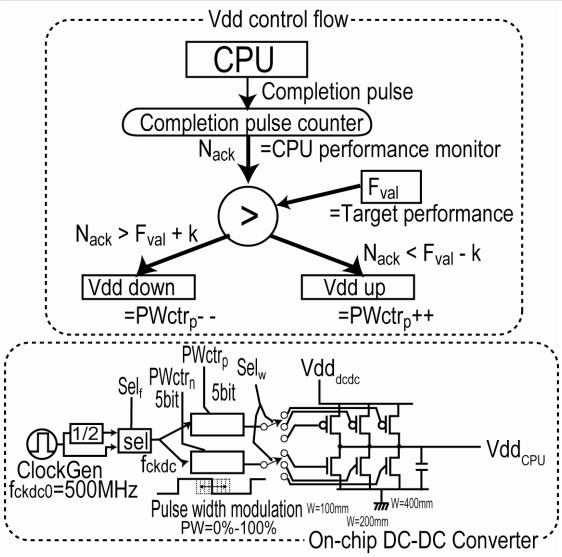

Figure 3.3.5: Power supply control flow and designed on-chip DC-DC converter.

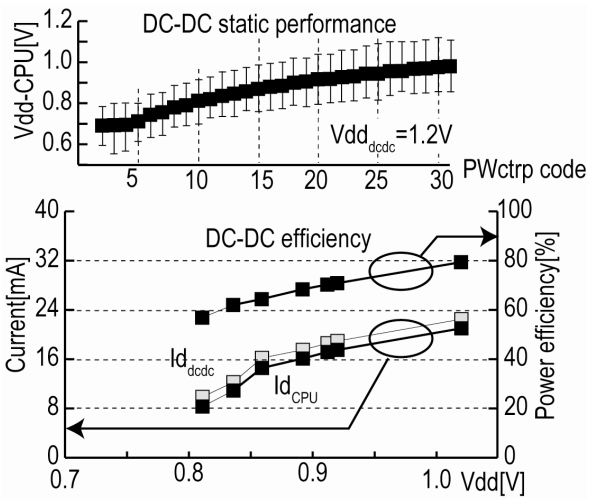

Figure 3.3.6: Measured results of on-chip DC-DC converter.

Continued on Page 587

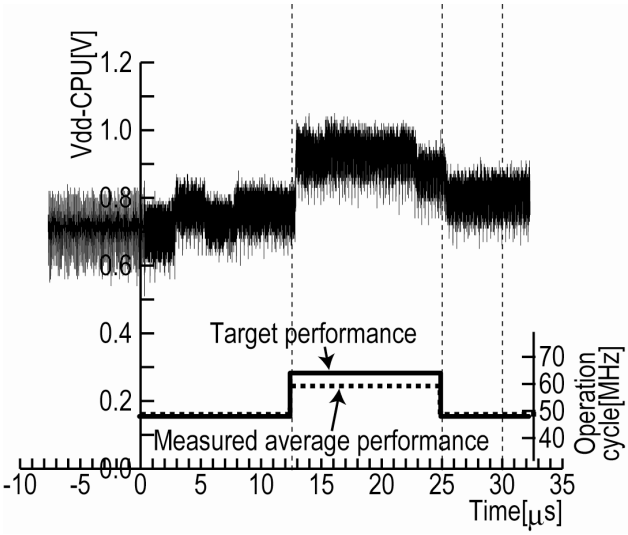

Figure 3.3.7: Measured results of target performance tracking.

